Cochrane Library

Cochrane Database of Systematic Reviews

AD-8 for diagnosis of dementia across a variety of healthcare settings (Protocol)

Hendry K, Lees RA, McShane R, Noel-Storr AH, Stott DJ, Quinn TJ

Hendry K, Lees RA, McShane R, Noel-Storr AH, Stott DJ, Quinn TJ.
AD-8 for diagnosis of dementia across a variety of healthcare settings.
Cochrane Database of Systematic Reviews 2014, Issue 5. Art. No.: CD011121.
DOI: 10.1002/14651858.CD011121.

www.cochranelibrary.com

WILEY

TABLE OF CONTENTS

[Diagnostic Test Accuracy Protocol]

AD-8 for diagnosis of dementia across a variety of healthcare settings

Kirsty Hendry[1], Rosalind A Lees[1], Rupert McShane[2], Anna H Noel-Storr[2], David J Stott[1], Terry J Quinn[1]

[1]Institute of Cardiovascular and Medical Sciences, University of Glasgow, Glasgow, UK. [2]Radcliffe Department of Medicine, University of Oxford, Oxford, UK

Contact address: Terry J Quinn, Institute of Cardiovascular and Medical Sciences, University of Glasgow, New Lister Building, Glasgow Royal Infirmary, Glasgow, G4 OSF, UK. Terry.Quinn@glasgow.ac.uk.

Editorial group: Cochrane Dementia and Cognitive Improvement Group
Publication status and date: New, published in Issue 5, 2014.

Citation: Hendry K, Lees RA, McShane R, Noel-Storr AH, Stott DJ, Quinn TJ. AD-8 for diagnosis of dementia across a variety of healthcare settings. *Cochrane Database of Systematic Reviews* 2014, Issue 5. Art. No.: CD011121. DOI: 10.1002/14651858.CD011121.

ABSTRACT

This is a protocol for a Cochrane Review (Diagnostic test accuracy). The objectives are as follows:

To determine the diagnostic accuracy of the informant-based questionnaire AD-8, in detection of all-cause (undifferentiated) dementia in adults. We will present data for each healthcare setting where AD-8 may be employed (community; primary care; secondary care).

BACKGROUND

Dementia is a substantial and growing public health concern (Hebert 2013; Prince 2013). As an example, depending on case definition employed, contemporary estimates of dementia prevalence in the United States are in the range 2.5 to 4.5 million individuals (1.6% population at higher rate; 6.5% of those aged over 65). Changes in population demographics will be accompanied by increases in global dementia incidence and prevalence (Ferri 2005). Although the magnitude of the increase in prevalent dementia may have been overestimated in previous prediction models (Matthews 2013), there is no doubt that absolute numbers of older adults with dementia will increase substantially in the short- to medium-term future (Ferri 2005).

A key element of effective management in dementia is a firm diagnosis. Recent guidelines place emphasis on early diagnosis to facilitate improved management and to allow informed discussions and planning with patients and carers. The benefits of screening for cognitive decline are debated (Brunet 2012); however in certain healthcare systems screening or case-finding has already been introduced for certain groups, e.g. unscheduled hospital admissions of older adults (Shenkin 2014).

Given the projected global increase in dementia prevalence, there is a potential tension between the clinical requirements for robust diagnosis at the individual patient level and the need for equitable, easy access to diagnosis at a population level. The ideal would be expert, multidisciplinary assessment informed by various supplementary investigations (neuropsychology; neuroimaging or other biomarkers). This approach is only really feasible in a specialist memory service and is not suited to population screening or case-finding.

In practice a two-stage process is often employed - with initial 'triage' assessments, suitable for use by non-specialists - to select those who require second-stage, further detailed assessment (Boustani 2003).

Various tools for initial cognitive screening have been described (Brodaty 2002; Folstein 1975; Galvin 2005). Regardless of the methods employed, there is scope for improvement, with observational work suggesting that many with dementia are not diagnosed (Chodosh 2004; Valcour 2000).

Screening assessment often takes the form of brief, direct cognitive testing. Such an approach will only provide a 'snapshot' of cognitive function. However, a defining feature of dementia is cognitive or neuropsychological change over time. People with cognitive problems themselves may struggle to make an objective assessment of personal change and so an attractive approach is to question collateral sources with sufficient knowledge of the person. These informant-based interviews aim to retrospectively assess change in function.

An instrument prevalent in research and clinical practice, particularly in North America, is the eight-item Informant Interview to differentiate Ageing and Dementia (AD-8) and this screening/triage tool will be the focus of this review.

A number of properties can be described for a clinical assessment (reliability, responsiveness, feasibility). For our purposes the test property of greatest interest is diagnostic test accuracy (DTA) (Cordell 2013).

Although we will describe test accuracy of AD-8 for dementia *diagnosis*, AD-8 used in isolation is not suitable for establishing a clinical dementia diagnosis. AD-8 is a triage tool, suitable for selecting those who require more definitive assessment.

Target condition being diagnosed

The target condition for this diagnostic test accuracy review is all-cause dementia (clinical diagnosis) (Appendix 1).

Dementia is a syndrome characterised by cognitive or neuropsychological decline sufficient to interfere with usual functioning. The neurodegeneration and clinical manifestations of dementia are progressive and at present there is no 'cure', although numerous pharmacological (Birks 2006; McShane 2006) and non-pharmacological interventions (Bahar-Fuchs 2013) to slow or arrest cognitive decline have been described.

Dementia remains a clinical diagnosis, based on history from the person and suitable collateral sources and direct examination, including cognitive assessment. There is no universally-accepted, ante-mortem, gold-standard diagnostic strategy. We have chosen expert clinical diagnosis as our gold standard (reference standard) for describing AD-8 test properties, as we believe this is most in keeping with current diagnostic criteria and best practice.

Dementia diagnosis can be made according to various internationally-accepted diagnostic criteria, with exemplars being the World Health Organization, International Classification of Diseases (ICD) ICD-10 and the American Psychiatric Association Diagnostic and Statistical Manual of Mental Disorders (DSM) DSM-IV for all-cause dementia and subtypes (Appendix 1). The label of dementia encompasses varying pathologies, of which Alzheimer's disease is the most common. Diagnostic criteria are available for specific dementia subtypes, i.e. NINCDS-ADRDA criteria for Alzheimer's dementia (McKhann 1984; McKhann 2011); McKeith criteria for Lewy Body dementia (McKeith 2005); Lund criteria for frontotemporal dementias (McKhann 2001); and the NINDS-AIREN criteria for vascular dementia (Román 1993).

Index test(s)

Our index test will be the Alzheimer Disease 8 (AD-8) (Galvin 2005) (Appendix 2; Appendix 3).

First published in 2005, the AD-8 is a screening tool which has been used to distinguish individuals with normal cognitive function from those with dementia or mild cognitive impairment. It is designed to be administered to a relevant proxy, usually a relative or carer, in questionnaire form. The AD-8 is a brief screening tool. With only eight questions it takes less than three minutes to complete and was developed to replace other lengthy informant questionnaires (Galvin 2006). The AD-8 was originally developed for administration in the English language but has been reproduced in other languages including Brazilian Portuguese (Correia 2011),Taiwanese (Yang 2011), and Korean (Ryu 2009).

The AD-8 items cover domains of judgement, hobby/activity level, repetitive conversations, learning ability, memory in relation to date/appointments, finances and daily thought processes. Informants indicate presence of change "over several years" using re-

sponses of 'Yes, a change', 'No, no change' or 'NA, don't know'. Each 'yes' answer is scored one point, giving scores ranging from zero, where no change has been noticed by the informant, to eight, where change has been noted across all domains. The commonly employed threshold score for AD-8 to differentiate cognitive from no cognitive impairment is greater than or equal to two out of eight (i.e. a 'yes' response for two or more items) (Galvin 2005).

AD-8 has a number of features that make it attractive for clinical and research use. The questions used have an immediacy and relevance that is likely to appeal to users. Assessment and (informant) scoring is brief, and as the scale is not typically interviewer-administered it requires minimal training in application and scoring. There are data to suggest that, compared to standard direct assessments, informant interviews may be less prone to bias from cultural norms and previous level of education (Jorm 2004). New diagnostic criteria for dementia make explicit reference to documenting decline and involving collateral informants, emphasising the potential utility of an informant interview tool such as AD-8.

Clinical pathway

Dementia develops over a trajectory of several years and screening tests may be performed at different stages in the dementia pathway. In this review we will consider any use of AD-8 as an initial assessment for cognitive decline and we will not limit to a particular healthcare setting. We have operationalised the various settings where AD-8 may be used as secondary care, primary care, and community.

In secondary-care settings, patients will have been referred for expert input but not exclusively due to memory complaints. Cognitive testing in secondary care involves two main groups: opportunistic screening of adults presenting as unscheduled admissions to hospitals, and those people referred to specific dementia, memory or psychiatry of older age services. Both populations will have a high prevalence of cognitive disorders and mimics. Secondary-care patients are more likely to have had a degree of prior cognitive assessment than those in other settings, although we recognise that cognitive screening prior to referral to specialist services is neither consistent nor guaranteed (Menon 2011).

In the general practice/primary care setting, the person will self present to a non-specialist service because of subjective memory complaints. There is unlikely to have been previous cognitive testing but prevalence of disease may be reasonably high. Using AD-8 in this setting could be described as 'triage' or 'case-finding'.

In the community setting, the cohort is largely unselected and the approach may be described as 'population screening'.

Most studies of test accuracy compare the test against contemporaneous reference standards (in this case, clinical dementia diagnosis). An alternative is to describe the test properties for detection of early, 'pre-clinical' problems that are formally diagnosed during prospective, longitudinal follow-up. This delayed verification approach is commonly employed in studies describing properties of dementia biomarkers, but may have utility for other test strategies such as informant interview.

Rationale

There is no consensus on the optimal initial assessment for dementia, and choice is currently dictated by experience with a particular instrument, time constraints and training. A better understanding of the diagnostic properties of various strategies would allow for an informed approach to testing. Critical evaluation of the evidence base for screening tests or other diagnostic markers is of major importance. Without a robust synthesis of the available information there is the risk that future research, clinical practice and policy will be built on erroneous assumptions about diagnostic validity.

AD-8 is commonly used in practice and research; it is used internationally and is one of only a few validated informant-based screening/diagnostic tools. A literature describing the test accuracy of AD-8 in different settings is available, although some of these studies have been modest in size. Thus systematic review and, if possible, meta-analysis of the diagnostic properties of AD-8 is warranted.

This review will form part of a body of work describing the diagnostic properties of commonly-used dementia tools (Appendix 4). At present we are conducting single-test reviews and meta-analyses. However the intention is then to collate these data, performing an overview allowing comparison of various test strategies.

OBJECTIVES

To determine the diagnostic accuracy of the informant-based questionnaire AD-8, in detection of all-cause (undifferentiated) dementia in adults. We will present data for each healthcare setting where AD-8 may be employed (community; primary care; secondary care).

Secondary objectives

Where data are available, we will describe the following:

1. Accuracy of AD-8 for early detection of cognitive problems, where a later diagnosis of dementia is made (delayed verification diagnosis).

2. Effects of heterogeneity on the reported diagnostic accuracy of AD-8. Potential sources of heterogeneity that we will explore include: case mix of cohort; method of dementia diagnosis; method AD-8 assessment.

METHODS

Criteria for considering studies for this review

Types of studies

We anticipate that the majority of studies will be of AD-8 test properties compared against a contemporaneous clinical diagnosis of dementia in secondary care settings. We will include test studies performed in other healthcare settings and classify these as: 'primary care' or 'community'. We will include studies that use a delayed verification methodology but will perform a separate analysis for this study design.

Case-control studies are known to potentially overestimate properties of a test and we will not include such studies in this review.

We will not include case studies or samples with very small numbers (chosen as 10 participants, for the purposes of this review).

There may be cases where settings are mixed, for example, a population study 'enriched' with additional cases from primary care. We will consider separate data from each setting if available. If these

data are not available we will treat these studies as case-control and will exclude them.

Participants

All adults (aged over 18 years) will be eligible.

We have not predefined exclusion criteria relating to the 'case-mix' of the population studied, but will assess this aspect of the study as part of our assessment of heterogeneity. Where there is concern that the participants are not representative of a primary care sample, we will explore this at study level using the 'Risk of bias' assessment framework outlined below.

Index tests

Studies must include (not necessarily exclusively) AD-8 used as an informant questionnaire.

AD-8 has been translated into various languages to allow international administration. The properties of a translated AD-8 in a cohort of non-English speakers may differ from properties of the original English language questionnaire. We will collect data on the principal language used for AD-8 assessment.

For this review we will not consider other cognitive screening/assessment tools. Where a paper describes AD-8 with in-study comparison against another screening tool, we will include the AD-8 data only. Where AD-8 is used in combination with another cognitive screening tool, we will include the AD-8 data only.

Target conditions

We will include any clinical diagnosis of all-cause (unspecified) dementia. We will not require defining of a particular dementia subtype, although where available we will record these data.

Reference standards

Our reference standard will be clinical diagnosis of dementia. We recognise that clinical diagnosis itself has a degree of variability but this is not unique to dementia studies and does not invalidate the basic diagnostic test accuracy approach.

Primary analysis will be for clinical diagnosis to include all-cause (unspecified) dementia, using any recognised diagnostic criteria (for example, ICD-10; DSM-IV ICD-10 DSM-IV). Dementia diagnosis may specify a pathological subtype and we will include all common dementia subtypes (e.g. NINCDS-ADRDA (Alzheimer's Disease), Lund-Manchester (Frontotemporal dementia), McKeith (Dementia with Lewy Bodies), NINCDS-AIREN (Vascular dementia); McKeith 2005; McKhann 1984; McKhann 2001; Román 1993). We have not defined preferred diagnostic criteria for rarer forms of dementia (e.g. alcohol-related; HIV-related; prion disease-related). We will consider these under our rubric of 'all-cause' dementia, and will not consider them separately.

The label 'dementia' can also span a range of disease severities, from mild disease to 'end stage'. The diagnostic properties of a tool will vary depending on disease stage, for example, true positives are more likely when disease is advanced and diagnosis is clear. For our primary analysis we will include any dementia diagnosis at any stage of disease.

Clinicians may use imaging, pathology or other data to aid diagnosis; however, we will not include diagnosis based only on these data without corresponding clinical assessment. We recognise that different iterations of diagnostic criteria may not be directly comparable and that diagnosis may vary with the degree or manner in which the criteria have been operationalised (e.g. individual clinician versus algorithm versus consensus determination). We will collect data on method and application of dementia diagnosis for each study, and will explore potential effects as part of our assessment of heterogeneity. We will not accept use of other (brief) direct performance tests in isolation as a basis for diagnosis.

We recognise that dementia diagnosis often comprises a degree of informant assessment. Thus there is potential for incorporation bias. We will explore the potential effects of this bias through our 'Risk of bias' assessment.

Search methods for identification of studies

We will use a variety of information sources to ensure all relevant studies are included. We will devise terms for electronic database searching in conjunction with the team at the Cochrane Dementia and Cognitive Improvement Group. As this AD-8 review forms part of a suite of reviews looking at informant scales we have created a comprehensive search strategy designed to pick up all cognitive assessment scales Quinn 2014; we will complement this generic search with searches specific to AD-8 terminology.

Electronic searches

We will search the specialised register of the Cochrane Dementia and Cognitive Improvement Group, ALOIS (which includes both intervention and diagnostic accuracy studies), MEDLINE (Ovid SP), EMBASE (Ovid SP), BIOSIS (Web of Knowledge), Science Citation Index (ISI Web of Knowledge), PsycINFO (Ovid SP), CINAHL (EBSCOhost) and LILACS (Bireme). See Appendix 5; Appendix 6 for a proposed draft strategy to be run in MEDLINE (Ovid SP) along with a narrative describing how the strategy was developed and validated. We will design similarly-structured search strategies using search terms appropriate to each database. We will use MeSH words and other controlled vocabulary where appropriate.

We will also search sources specific to diagnostic accuracy or systematic review:

· MEDION database (Meta-analyses van Diagnostisch Onderzoek www.mediondatabase.nl);

· DARE (Database of Abstracts of Reviews of Effects) and HTA Database (Health Technology Assessments Database), both *The Cochrane Library*;

· ARIF database (Aggressive Research Intelligence Facility; www.arif.bham.ac.uk).

We will apply no language or date restrictions to the electronic searches and will operate no restrictions by publication status; assessing for potential inclusion: abstracts, conference proceedings and unpublished data. We will use translation services as necessary.

The Cochrane Dementia and Cognitive Impairment Group Trials Search Co-ordinator will run the initial searches.

Searching other resources

Grey literature and proceedings: chosen electronic databases include assessments of conference proceedings. We will aim to access theses or PhD abstracts from institutions known to be involved in prospective dementia studies.

Handsearching: We will not perform handsearching as there is little published evidence of the benefits of handsearching for diagnostic studies (Glanville 2010).

Reference lists: We will check the reference lists of all relevant studies and reviews in the field for further possible titles and will repeat the process until no new titles are found (Greenhalgh 2005).

Correspondence: We will contact research groups who have published or are conducting work on AD-8 for dementia diagnosis, informed by results of initial search.

We will use relevant studies in PubMed to search for additional studies using the 'Related article' feature. We will examine key studies in citation databases such as Science Citation Index and Scopus to ascertain any further relevant studies.

Data collection and analysis

Selection of studies

Two review authors will independently screen all titles generated by electronic database searches for relevance. Two review authors will inspect abstracts of selected titles and will select all potentially eligible studies for full-paper review. Two review authors will independently assess full manuscripts against the inclusion criteria, resolving disagreement by discussion, or by involving an arbitrator if necessary.

Where a study may include useable data but these are not presented in the published manuscript, we will contact the authors directly to request further information. If the same data are presented in more than one paper we will include the primary paper only.

We will detail the study selection process in a PRISMA flow diagram.

Data extraction and management

We will extract data to a study-specific pro forma that includes clinical/demographic details of the participants; details of setting; details of AD-8 administration and details of the dementia diagnosis process.

We will extract test accuracy data to a standard two-by-two table.

Two review authors, blinded to study identifiers, will perform data extraction independently, resolving disagreement by discussion, with the use of an arbitrator if necessary.

For each included paper, we will detail the flow of participants (numbers recruited, included, assessed) in a flow diagram.

Assessment of methodological quality

As well as describing test accuracy, an important goal of the DTA process is to improve study design and reporting in dementia diagnostic studies. For this reason we will assess methodological and reporting quality using two complementary processes.

We will assess the quality of study reporting using the STARD checklist (Bossuyt 2003) (Appendix 7). If it becomes available during the course of the review, we will use the proposed dementia-specific extension to the STARD tool - STARDdem (starddem.org/). We will tabulate and present STARD data as an appendix to the review.

We will assess the methodological quality of each study using the QUADAS-2 tool (www.bris.ac.uk/quadas/quadas-2).(Appendix 8) This tool incorporates domains specific to participant selection; index test; reference standard; and participant flow. Each domain is assessed for risk of bias and the first three domains are also assessed for applicability. Certain key areas important for quality assessment are participant selection; blinding; and missing data. Following a group meeting of review authors, we created guidance for the application of QUADAS-2 to dementia screening assessments, specifically developing anchoring statements for QUADAS-based assessment that are suited to dementia test accuracy studies. This QUADAS-2 guidance was created through a multidisciplinary working group and has been extensively piloted. The process and resulting statements for assessment are described in Appendix 9.

We will not use QUADAS-2 data to form a summary quality score, but will produce a narrative summary describing numbers of studies that found high/low/unclear risk of bias/concerns regarding applicability with corresponding tabular and graphical displays.

Paired independent raters, blinded to each other's scores, will perform both assessments, resolving disagreement by further review and discussion, with recourse to a third party arbitrator where necessary.

Statistical analysis and data synthesis

We are interested in the test accuracy of AD-8 for the dichotomous variable 'dementia'/'no dementia'. Thus, we will apply the current Cochrane DTA framework for analysis of a single test. We will extract data from included papers to allow creation of a standard two-by-two data table showing dichotomised AD-8 test results (AD-8 positive or AD-8 negative) cross-classified with binary reference standard (dementia or no dementia).

We will use Review Manager 5 software (RevMan 2012) to calculate sensitivity, specificity and 95% confidence intervals (CIs) from the two-by-two tables abstracted from the included studies. We will use a threshold score of two or more on AD-8 for primary analyses. If data at other thresholds are presented we will examine these in separate analyses. We will present individual study results graphically by plotting estimates of sensitivities and specificities as forest plots.

To allow for summary analysis, we will use software additional to RevMan (SAS release 9.1). Using the bivariate approach we will describe metrics of pooled sensitivity, specificity, positive and negative predictive values; positive and negative likelihood ratios, all with corresponding 95% confidence intervals. We will plot summary data in receiver operating characteristic (ROC) space, including 95% confidence and prediction regions.

We suspect papers will use the classical cross-sectional test accuracy study design. An alternative is the 'delayed verification' study design, i.e. AD-8 is performed at baseline and those without disease are prospectively followed up for development of incident dementia. If any papers use this methodology, we will use baseline (con-

temporaneous testing) data for our primary analysis; we will not describe data based on prospective testing separately.

Investigations of heterogeneity

Heterogeneity is expected in DTA reviews and 'traditional' measures of heterogeneity used in meta-analysis are not appropriate to DTA reviews.

The properties of a tool describe behaviour of the instrument under particular circumstances. We will include all AD-8 studies in narrative review. We have prespecified particular areas of potential heterogeneity.

Healthcare setting

We suspect that healthcare setting will impact on properties of the test and we will restrict our primary analyses to the various predefined healthcare settings. Based on the spread of data, review authors will decide if a pooled analysis including all healthcare settings is valid. If we conduct it, we will clearly label this analysis as a secondary analysis.

Case mix

In the first instance we will explore age, taking age over 65 years as a reference point. We suspect that the majority of included participants in eligible studies will be aged over 65 years. AD-8 may have different properties in younger cohorts and so we will look at age ranges within studies. We will grade studies that have greater than 20% of included participants younger than 65 years as potentially unrepresentative, and will run sensitivity analyses assessing the effect on summary parameters of removing these studies.

We anticipate that most studies will be of unselected adults. However, studies may limit inclusion to a specific population, for example, stroke survivors. Again, we will run sensitivity analyses assessing the effect on summary parameters of removing these studies.

If data allow, we will also explore case mix within the dementia diagnosis, performing subgroup or sensitivity analyses for various dementia pathological subtypes.

Clinical criteria used to reach dementia diagnosis

We will record the classification used (for example ICD-10; DSM-IV ICD-10 DSM-IV) and the methodology used to reach dementia diagnosis (for example, individual assessment; group (consensus) assessment). We will assess the 'quality' of diagnosis at study level using the QUADAS-2 tool. If data allow, we will compare subgroups, describing summary statistics for each group, for example ICD-based diagnosis versus DSM-based diagnosis; individual clinician versus group adjudication.

Technical features of the testing strategy

Our focus will be on language of assessment. In the first instance we will classify as English language and non-English language tests. We will perform subgroup analyses comparing English language AD-8 versus non-English language AD-8.

We do not anticipate that AD-8 will be used at differing thresholds. If papers describe a range of AD-8 cut-points we will explore this using HSROC techniques, with cut-points as covariates.

Sensitivity analyses

Where appropriate (i.e. if not already explored in our analyses of heterogeneity) and as data allow, we will explore the sensitivity of any summary accuracy estimates to aspects of study quality such as nature of blinding and loss to follow-up, guided by the anchoring statements developed in our QUADAS-2 exercise. Primary analysis will include all eligible studies, while sensitivity analysis will exclude studies of low quality (high likelihood of bias) to determine if the results are influenced by inclusion of the lower-quality studies.

Assessment of reporting bias

We will not investigate reporting bias because of current uncertainty about how it operates in test accuracy studies, and about the interpretation of existing analytical tools such as the funnel plot.

REFERENCES

Additional references

Bahar-Fuchs 2013

Bahar-Fuchs A, Clare L, Woods B. Cognitive training and cognitive rehabilitation for mild to moderate Alzheimer's disease and vascular dementia. *Cochrane Database of Systematic Reviews* 2013, Issue 6. [DOI: 10.1002/14651858.CD003260.pub2]

Birks 2006

Birks J. Cholinesterase inhibitors for Alzheimer's disease. *Cochrane Database of Systematic Reviews* 2006, Issue 1. [DOI: 10.1002/14651858.CD005593]

Bossuyt 2003

Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, Irwig LM, et al. Standards for Reporting of Diagnostic Accuracy. Towards complete and accurate reporting of studies of diagnostic accuracy: The STARD initiative. *BMJ* 2003;**326**(7379):41-4.

Boustani 2003

Boustani M, Peterson B, Hanson L, Harris R, Lohr KN. Screening for dementia in primary care: a summary of the evidence for the US Preventative Services Task Force. *Annals of Internal Medicine* 2003;**138**(11):927-37.

Brodaty 2002

Brodaty H. The GPCOG: a new screening test for dementia designed for general practice. *Journal of the American Geriatrics Society* 2002;**50**(3):530-4.

Brunet 2012

Brunet MD, McCartney M, Heath I, Tomlinson J, Gorodon P, Cosgrove J, et al. There is no evidence base for proposed dementia screening. *BMJ* 2012;**345**:e8588.

Chodosh 2004

Chodosh J, Petitti DB, Elliott M, Hays RD, Crooks VC, Reuben DB, et al. Physician recognition of cognitive impairment: evaluating the need for improvement. *Journal of the American Geriatrics Society* 2004;**52**(7):1051-9.

Cordell 2013

Cordell CB, Borson S, Boustani M, Chodosh J, Reuben D, Verghese J, et al. Medicare Detection of Cognitive Impairment Workgroup. Alzheimer's Association recommendations for operationalizing the detection of cognitive impairment during Medicare Annual Wellness Visit in a primary care setting. *Alzheimer's & Dementia* 2013;**9**(2):141-50.

Correia 2011

Correia CC, Lima F, Junqueira F, Campos MS, Bastos O, Petribu K, et al. AD8-Brazil: cross-cultural validation of the ascertaining dementia interview in Portuguese. *Journal of Alzheimer's Disease* 2011;**27**(1):177-85.

Davis 2013

David HJD, Creavin ST, Noel-Storr A, Quinn TJ, Smailagic N, Hyde C, Brayne C, McShane R, Cullum S. Neuropsychological tests for the diagnosis of Alzheimer's disease dementia and other dementias: a generic protocol for cross-sectional and delayed verification studies.. *Cochrane Database of Systematic Reviews* 2013, Issue 4; CD010460. [DOI: 10.1002/14651858.CD010460]

DSM-IV

American Psychiatric Association. Diagnostic and statistical manual of mental disorders (4th Ed). Washington DC: American Psychiatric Association, 2000.

Ferri 2005

Ferri CP, Prince M, Brayne C, Brodaty H, Fratiglioni L, Ganguli M, et al. Alzheimer's Disease International. Global prevalence of dementia: a Delphi consensus study. *Lancet* 2005;**366**(9503):2112-7.

Folstein 1975

Folstein MF, Folstein SE, McHugh PR. "Minimental state": a practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research* 1975;**12**(3):189-98.

Galvin 2005

Galvin JE, Roe CM, Powlishta KK, Coats MA, Muich SJ, Grant E, et al. The AD8: A brief informant interview to detect dementia. *Neurology* 2005;**65**(4):559-64.

Galvin 2006

Galvin JE, Roe CM, Xiong C, Morris JC. Validity and reliability of the AD8 informant interview in dementia.. *Neurology* 2006;**67**(11):1942-8.

Glanville 2010

Glanville JM, Cikalo M, Crawford F, Dozier M, Lowson P. Handsearching for reports of diagnostic test accuracy studies adding to the evidence base.. Joint Cochrane and Campbell Colloquium, Madrid, Spain. 2010.

Greenhalgh 2005

Greenhalgh T, Peacock R. Effectiveness and efficiency of search methods in systematic reviews of complex evidence. *BMJ* 2005;**331**(7524):1064-5.

Hebert 2013

Hebert LE, Weuve J, Scherr PA, Evans DA. Alzheimer disease in the United States (2010-2050) estimated using the 2010 census. *Neurology* 2013;**80**(19):1778-83.

ICD-10

World Health Organisation. International Classification of Diseases (ICD) 10th Edition. WHO, 1992.

Jorm 2004

Jorm AF. The Informant Questionnaire on cognitive decline in the elderly (IQCODE): a review. *International Psychogeriatrics* 2004;**16**(3):275-93.

Matthews 2013

Matthews FE, Arthur A, Barnes LE, Bond J, Jagger C, Robinson L, et al. A two-decade comparison of prevalence of dementia in individuals aged 65 and older from three geographical areas of England: results of the Cognitive Function and Ageing Study I and II. *Lancet* 2013;**382**(9902):1405-12.

McKeith 2005

McKeith IG, Dickson DW, Lowe J, Emre M, O'Brien JT, Feldman H, et al. Diagnosis and management of dementia with Lewy bodies: third report of the DLB Consortium. *Neurology* 2005;**65**(12):1863-72.

McKhann 1984

McKhann G, Drachman D, Folstein M, Katzman R, Price D, Stadlan EM. Clinical diagnosis of Alzheimer's disease: report of the NINCDS-ADRDA Work Group under the auspices of Department of Health and Human Services Task Force on Alzheimer's Disease. *Neurology* 1984;**34**(7):939-44.

McKhann 2001

McKhann GM, Albert MS, Grossman M, Miller B, Dickson D, Trojanowski JQ: Work Group on Frontotemporal dementia and PIck's disease. Clinical and pathological diagnosis of frontotemporal dementia: report of the Work Group on Frontotemporal Dementia and Pick's Disease. *Archives of Neurology* 2001;**58**(11):1803-9.

McKhann 2011

McKhann GM, Knopman DS, Chertkow H, Hyman BT, Jack CR Jr, Kawas CH, et al. The diagnosis of dementia due to Alzheimer's disease:recommendations from the National Institute on Aging - Alzheimer's Association workgroup on diagnostic guidelines for Alzheimer's disease. *Alzheimer's & Dementia* 2011;**7**(3):263-9.

McShane 2006

McShane R, Areosa Sastre A, Minakaran N. Memantine for dementia. *Cochrane Database of Systematic Reviews* 2006, Issue 2. [DOI: 10.1002/14651858.CD003154]

Menon 2011

Menon R, Larner AJ. Use of cognitive screening instruments in primary care: the impact of national dementia directives. *Family Practice* 2011;**28**(3):272-6..

Prince 2013

Prince M, Bryce R, Albanese E, Wimo A, Ribeiro W, Ferri CP. The global prevalence of dementia: a systematic review and metaanalysis. *Alzheimer's and Dementia* 2013;**9**(1):63-75.

Quinn 2014

Quinn TJ, Fearon P, Noel-Starr A, Young C, McShane R, Stott DJ. Informant Questionnaire on Cognitive Decline in the Elderly (IQCODE) for the diagnosis of dementia within community dwelling populations. *Cochrane Database of Systematic Reviews* 2014, Issue 4. [DOI: 10.1002/14651858.CD010079.pub2]

RevMan 2012 [Computer program]

The Nordic Cochrane Centre, The Cochrane Collaboration. Review Manager (RevMan). Version 5.2. Copenhagen: The Nordic Cochrane Centre, The Cochrane Collaboration, 2012.

Román 1993

Román GC, Tatemichi TK, Erkinjutti T, Cummings JL, Masdeu JC, Garcia JH, et al. Vascular dementia: diagnostic criteria for research studies. Report of the NINDS-AIREN International Workshop. *Neurology* 1993;**43**(2):250-60.

Ryu 2009

Ryu HJ, Kim HJ, Han SH. Validity and reliability of the Korean version of AD8 informant interview (K-AD8) in dementia. *Alzheimer Disease and Associated Disorders* 2009;**23**(4):371-6.

Shenkin 2014

Shenkin SD, Russ TC, Ryan TM, MacLullich AM. Screening for dementia and other causes of cognitive impairment in general hospital inpatients. *Age and Ageing* 2014;**43**(2):166-8.

Valcour 2000

Valcour VG, Masaki KH, Curb JD, Blanchette PL. The detection of dementia in the primary care setting. *Archives of Internal Medicine* 2000;**160**(19):2964-8.

Yang 2011

Yang YH, Galvin JE, Morris JC, Lai CL, Chou MC, Liu CK. Application of AD8 questionnaire to screen very mild dementia in Taiwanese. *American Journal of Alzheimers Disease and Other Dementias* 2011;**26**(2):134-8.

APPENDICES

Appendix 1. WHO International Classification of Disease - Dementia

World Health Organization International Classification of Diseases 10

F00 - F09 ORGANIC, INCLUDING SYMPTOMATIC, MENTAL DISORDERS

DEMENTIA

G1. Evidence of each of the following:

(1) A decline in memory, which is most evident in the learning of new information, although in more severe cases, the recall of previously learned information may be also affected. The impairment applies to both verbal and non-verbal material. The decline should be objectively verified by obtaining a reliable history from an informant, supplemented, if possible, by neuropsychological tests or quantified cognitive assessments. The severity of the decline, with mild impairment as the threshold for diagnosis, should be assessed as follows:

Mild: a degree of memory loss sufficient to interfere with everyday activities, though not so severe as to be incompatible with independent living. The main function affected is the learning of new material. For example, the individual has difficulty in registering, storing and recalling elements in daily living, such as where belongings have been put, social arrangements, or information recently imparted by family members.

Moderate: A degree of memory loss which represents a serious handicap to independent living. Only highly learned or very familiar material is retained. New information is retained only occasionally and very briefly. The individual is unable to recall basic information about where he lives, what he has recently been doing, or the names of familiar persons.

Severe: a degree of memory loss characterized by the complete inability to retain new information. Only fragments of previously learned information remain. The subject fails to recognize even close relatives.

(2) A decline in other cognitive abilities characterized by deterioration in judgement and thinking, such as planning and organizing, and in the general processing of information. Evidence for this should be obtained when possible from interviewing an informant, supplemented, if possible, by neuropsychological tests or quantified objective assessments. Deterioration from a previously higher level of performance should be established. The severity of the decline, with mild impairment as the threshold for diagnosis, should be assessed as follows:

Mild. The decline in cognitive abilities causes impaired performance in daily living, but not to a degree making the individual dependent on others. More complicated daily tasks or recreational activities cannot be undertaken.

Moderate. The decline in cognitive abilities makes the individual unable to function without the assistance of another in daily living, including shopping and handling money. Within the home, only simple chores are preserved. Activities are increasingly restricted and poorly sustained.

Severe. The decline is characterized by an absence, or virtual absence, of intelligible ideation. The overall severity of the dementia is best expressed as the level of decline in memory or other cognitive abilities, whichever is the more severe (e.g. mild decline in memory and moderate decline in cognitive abilities indicate a dementia of moderate severity).

G2. Preserved awareness of the environment during a period of time long enough to enable the unequivocal demonstration of G1. When there are superimposed episodes of delirium the diagnosis of dementia should be deferred.

G3. A decline in emotional control or motivation, or a change in social behaviour, manifest as at least one of the following:

(1) emotional lability;

(2) irritability;

(3) apathy;

(4) coarsening of social behaviour.

G4. For a confident clinical diagnosis, G1 should have been present for at least six months; if the period since the manifest onset is shorter, the diagnosis can only be tentative.

Comments: The diagnosis is further supported by evidence of damage to other higher cortical functions, such as aphasia, agnosia, apraxia.

Judgment about independent living or the development of dependence (upon others) need to take account of the cultural expectation and context.

Dementia is specified here as having a minimum duration of six months to avoid confusion with reversible states with identical behavioural syndromes, such as traumatic subdural haemorrhage (S06.5), normal pressure hydrocephalus (G91.2) and diffuse or focal brain injury (S06.2 and S06.3).

A fifth character may be used to indicate the presence of additional symptoms, in the categories F00-F03

(F00 Dementia in Alzheimer's disease; F01 Vascular dementia; F02 Dementia in diseases classified elsewhere; and

F03 Unspecified dementia), as follows:

.x0 without additional symptoms

.x1 with other symptoms, predominantly delusional

.x2 with other symptoms, predominantly hallucinatory

.x3 with other symptoms, predominantly depressive

.x4 with other mixed symptoms

A sixth character may be used to indicate the severity of the dementia:

.xx0 mild

.xx1 moderate

.xx2 severe

As mentioned above the overall severity of the dementia depends on the level of memory or intellectual impairment, whichever is the more severe.

F00 DEMENTIA IN ALZHEIMER'S DISEASE

A. The general criteria for dementia (G1 to G4) must be met.

B. There is no evidence from the history, physical examination or special investigations for any other possible cause of dementia (e.g. cerebrovascular disease, Parkinson's disease, Huntington's disease, normal pressure hydrocephalus), a systemic disorder (e.g. hypothyroidism, vit. B12 or folic acid deficiency, hypercalcaemia), or alcohol- or drug-abuse.

Comments: The diagnosis is confirmed by post mortem evidence of neurofibrillary tangles and neuritic plaques in excess of those found in normal ageing of the brain.

The following features support the diagnosis, but are not necessary elements: Involvement of cortical functions as evidenced by aphasia, agnosia or apraxia; decrease of motivation and drive, leading to apathy and lack of spontaneity; irritability and disinhibition of social behaviour; evidence from special investigations that there is cerebral atrophy, particularly if this can be shown to be increasing over time. In severe cases there may be Parkinson-like extrapyramidal changes, logoclonia, and epileptic fits.

Specification of features for possible subtypes. Because of the possibility that subtypes exist, it is recommended that the following characteristics be ascertained as a basis for a further classification: age at onset; rate of progression; the configuration of the clinical features, particularly the relative prominence (or lack) of temporal, parietal or frontal lobe signs; any neuropathological or neurochemical abnormalities, and their pattern.

The division of AD into subtypes can at present be accomplished in two ways: first by taking only the age of onset and labelling AD as either early or late, with an approximate cut-off point at 65 years; or secondly, by assessing how well the individual conforms to one of the two putative syndromes, early or late onset type. It should be noted that it is unlikely that a sharp distinction exists between early and late onset type. Early onset type may occur in late life, just as late onset type may occasionally have an onset under the age of 65. The following criteria may be used to differentiate F00.0 from F00.1, but it should be remembered that the status of this subdivision is still controversial.

F00.0 Dementia in Alzheimer's disease with early onset

1. The criteria for dementia in Alzheimer's disease (F00) must be met, and the age at onset being under 65 years.

2. In addition, at least one of the following requirements must be met:

(a) evidence of a relatively rapid onset and progression;

(b) in addition to memory impairment, there is aphasia (amnesic or sensory), agraphia, alexia, acalculia, or apraxia (indicating the presence of temporal, parietal and/or frontal lobe involvement).

F00.1 Dementia in Alzheimer's disease with late onset

1. The criteria for dementia in Alzheimer's disease (F00) must be met and the age at onset must be 65 or more.

2. In addition, at least one of the following requirements must be met:

(a) evidence of a very slow, gradual onset and progression (the rate of the latter may be known only retrospectively after a course of 3 years or more);

(b) predominance of memory impairment G1.1, over intellectual impairment G1.2 (see general criteria for dementia).

F00.2 Dementia in Alzheimer's disease, atypical or mixed type

Use this term and code for dementias that have important atypical features or that fulfil criteria for both early and late onset type of Alzheimer's disease. Mixed Alzheimer's and vascular dementia is also included here.

F00.9 Dementia in Alzheimer's disease, unspecified

F01 VASCULAR DEMENTIA

G1. The general criteria for dementia (G1 to G4) must be met.

G2. Unequal distribution of deficits in higher cognitive functions, with some affected and others relatively spared. Thus memory may be quite markedly affected while thinking, reasoning and information processing may show only mild decline.

G3. There is clinical evidence of focal brain damage, manifest as at least one of the following:

(1) unilateral spastic weakness of the limbs;

(2) unilaterally increased tendon reflexes;

(3) an extensor plantar response;

(4) pseudobulbar palsy.

G4. There is evidence from the history, examination, or tests, of a significant cerebrovascular disease, which may reasonably be judged to be etiologically related to the dementia (e.g. a history of stroke; evidence of cerebral infarction).

The following criteria may be used to differentiate subtypes of vascular dementia, but it should be remembered that the usefulness of this subdivision may not be generally accepted.

F01.0 Vascular dementia of acute onset

A. The general criteria for vascular dementia (F01) must be met.

B. The dementia develops rapidly (i.e. usually within one month, but within no longer than three months) after a succession of strokes, or (rarely) after a single large infarction.

F01.1 Multi-infarct dementia

A. The general criteria for vascular dementia (F01) must be met.

B. The onset of the dementia is gradual (i.e. within three to six months), following a number of minor ischaemic episodes.

Comments: It is presumed that there is an accumulation of infarcts in the cerebral parenchyma. Between the ischaemic episodes there may be periods of actual clinical improvement.

F01.2 Subcortical vascular dementia

A. The general criteria for vascular dementia (F01) must be met.

B. A history of hypertension.

C. Evidence from clinical examination and special investigations of vascular disease located in the deep white matter of the cerebral hemispheres, with preservation of the cerebral cortex.

F01.3 Mixed cortical and subcortical vascular dementia

Mixed cortical and subcortical components of the vascular dementia may be suspected from the clinical features, the results of investigations (including autopsy), or both.

F01.8 Other vascular dementia

F01.9 Vascular dementia, unspecified

F02 DEMENTIA IN OTHER DISEASES CLASSIFIED ELSEWHERE

F02.0 Dementia in Pick's disease

A. The general criteria for dementia (G1 to G4) must be met.

B. Slow onset with steady deterioration.

C. Predominance of frontal lobe involvement evidenced by two or more of the following:

(1) emotional blunting;

(2) coarsening of social behaviour;

(3) disinhibition;

(4) apathy or restlessness;

(5) aphasia.

D. Relative preservation, in the early stages, of memory and parietal lobe functions.

F02.1 Dementia in Creutzfeldt-Jakob disease

A. The general criteria for dementia (G1 to G4) must be met.

B. Very rapid progression of the dementia, with disintegration of virtually all higher cerebral functions.

C. The emergence, usually after or simultaneously with the dementia, of one or more of the following types of neurological symptoms and signs:

(1) pyramidal symptoms;

(2) extrapyramidal symptoms;

(3) cerebellar symptoms;

(4) aphasia;

(5) visual impairment.

Comments: An akinetic and mute state is the typical terminal stage. An amyotrophic variant may be seen, where the neurological signs precede the onset of the dementia. A characteristic electroencephalogram (periodic spikes against a slow and low voltage background), if present in association with the above clinical signs, will increase the probability of the diagnosis. However, the diagnosis can be confirmed only by neuropathological examination (neuronal loss, astrocytosis, and spongiform changes). Because of the risk of infection, this should be carried out only under special protective conditions.

F02.2 Dementia in Huntington's disease

A. The general criteria for dementia (G1 to G4) must be met.

B. Subcortical functions are affected first and dominate the picture of dementia throughout; manifest as slowness of thinking or movement and personality alteration with apathy or depression.

C. Presence of involuntary choreiform movements, typically of the face, hands or shoulders, or in the gait. The patient may attempt to conceal them by converting them into a voluntary action.

D. A history of Huntington's disease in one parent or a sibling; or a family history which suggests the disorder.

E. The absence of clinical features otherwise accounting for the abnormal movements.

Comments: In addition to involuntary choreiform movements there may be development of extrapyramidal rigidity or spasticity with pyramidal signs.

F02.3 Dementia in Parkinson's disease

A. The general criteria for dementia (G1 to G4) must be met.

B. Diagnosis of Parkinson's disease.

C. Absence of cognitive impairment attributable to anti-parkinsonian medication.

D. There is no evidence from the history, physical examination or special investigations for any other possible cause of dementia, including other forms of brain disease, damage or dysfunction (e.g. cerebrovascular disease, HIV disease, Huntington's disease, normal pressure hydrocephalus), a systemic disorder (e.g. hypothyroidism, vit. B12 or folic acid deficiency, hypercalcaemia), or alcohol or drug abuse.

If criteria are also fulfilled for dementia in Alzheimer's disease with late onset (F00.1), this category F00.1 should be used in combination with Parkinson's disease G20.

F02.4 Dementia in human immunodeficiency (HIV) disease

A. The general criteria for dementia (G1 to G4) must be met.

B. Diagnosis of HIV infection.

C. There is no evidence from the history, physical examination or special investigations for any other possible cause of dementia, including other forms of brain disease, damage or dysfunction (e.g. Alzheimer's disease, cerebrovascular disease, Parkinson's disease, Huntington's disease, normal pressure hydrocephalus), a systemic disorder (e.g. hypothyroidism, vit. B12 or folic acid deficiency, hypercalcaemia), or alcohol or drug abuse.

F02.8 Dementia in other specified diseases classified elsewhere

Dementia can occur as a manifestation or consequence of a variety of cerebral and somatic conditions. To specify the etiology, the ICD-10 code for the underlying condition should be added.

F03 UNSPECIFIED DEMENTIA

This category should be used when the general criteria for dementia are met, but when it is not possible to identify one of the specific types (F00.0-F02.9).

Appendix 2. AD8 Informant Questionnaire

Remember, “Yes, a change” indicates that there has been a change in the last several years caused by cognitive (thinking and memory) problems.	**YES** **A change**	**NO** **No change**	**N/A** **Don’t know**
1. Problems with judgment (e.g., problems making decisions, bad financial decisions, problems with thinking)			
2. Less interest in hobbies/activities			
3. Repeats the same things over and over (questions, stories, or statements)			
4. Trouble learning how to use a tool, appliance, or gadget (e.g., VCR, computer, microwave, remote control)			
5. Forgets correct month or year			
6. Trouble handling complicated financial affairs (e.g. balancing checkbook, income taxes, paying bills)			
7. Trouble remembering appointments			
8. Daily problems with thinking and/or memory			
TOTAL AD8 SCORE			

Appendix 3. AD8 administration and scoring guidelines

A spontaneous self-correction is allowed for all responses without counting as an error.

The questions are given to the respondent on a clipboard for self–administration or can be read aloud to the respondent either in person or over the phone. It is preferable to administer the AD8 to an informant, if available. If an informant is not available, the AD8 may be administered to the patient.

When administered to an informant, specifically ask the respondent to rate change in the patient.

When administered to the patient, specifically ask the patient to rate changes in his/her ability for each of the items, ***without*** attributing causality.

If read aloud to the respondent, it is important for the clinician to carefully read the phrase as worded and give emphasis to note changes due to cognitive problems (not physical problems).

There should be a one second delay between individual items.

No timeframe for change is required.

The final score is a sum of the number items marked “Yes, A change”.

Appendix 4. Commonly used cognitive assessments/screening tools

TEST	Cochrane DTA review in process
Mini-mental state examination (MMSE)	YES
GPcog	YES
Minicog	YES
Memory Impairment Screen (MIS)	Still available
Abbreviated mental testing	Still available
Clock drawing tests (CDT)	Still available
Montreal Cognitive Assessment (MoCA)	YES
IQCODE (informant interview)	YES

For each test, the planned review will encompass diagnostic test accuracy in community; primary and secondary care settings.

Appendix 5. Search Strategy for use with MEDLINE electronic database

MEDLINE In-process and other non-indexed citations and MEDLINE 1950-present (Ovid SP)	1. AD8.ti,ab. 2. "informant questionnaire on cognitive decline".ti,ab. 3. "Alzheimer’s Disease eight question screen".ti,ab. 4. "AD 8".ti,ab. 5. ("informant* questionnair*" adj3 (dement* or screening)).ti,ab. 6. ("screening test*" adj2 (dement* or alzheimer*)).ti,ab. 7. or/1-6

Appendix 6. Search strategy (MEDLINE Ovid SP) run for specialised register (ALOIS)

Search narrative: The searches detailed above are very simple, essentially single concept strategies based on the index test (AD8). This is a sensitive approach to take. More complex and developed searches are run each month for the dementia group.

Every month the following strategy is run in Medline (via Ovid SP). The results are screened based on a reading of title and abstract. The full texts (where there is one) are then obtained and a few key details about each study are extracted including Index test/s and details of population and setting. For this review it was expected that most studies would be identified through a search of multiple sources based on one concept (the index test in question). However, we felt it was worth also searching ALOIS for any studies which had evaluated the accuracy of AD8 but had not referred to it in the bibliographic details of the reference.

MEDLINE In-process and other non-indexed citations and MEDLINE 1950-present (Ovid SP)	1. "word recall".ti,ab. 2. "7-minute screen".ti,ab. 3. "6 item cognitive impairment test".ti,ab. 4. "6 CIT".ti,ab. 5. "AB cognitive screen".ti,ab. 6. "abbreviated mental test".ti,ab. 7. "ADAS-cog".ti,ab. 8. AD8.ti,ab. 9. "inform* interview".ti,ab. 10. "animal fluency test".ti,ab. 11. "brief alzheimer* screen".ti,ab. 12. "brief cognitive scale".ti,ab. 13. "clinical dementia rating scale".ti,ab. 14. "clinical dementia test".ti,ab. 15. "community screening interview for dementia".ti,ab. 16. "cognitive abilities screening instrument".ti,ab. 17. "cognitive assessment screening test".ti,ab. 18. "cognitive capacity screening examination".ti,ab. 19. "clock drawing test".ti,ab. 20. "deterioration cognitive observee".ti,ab. 21. "Dem Tect".ti,ab. 22. "fuld object memory evaluation".ti,ab. 23. "IQCODE".ti,ab. 24. "mattis dementia rating scale".ti,ab. 25. "memory impairment screen".ti,ab. 26. "minnesota cognitive acuity screen".ti,ab. 27. "mini-cog".ti,ab. 28. "mini-mental state exam*".ti,ab. 29. "mmse".ti,ab. 30. "modified mini-mental state exam".ti,ab.

(Continued)

31. "3MS".ti,ab.
32. "neurobehavioural cognitive status exam*".ti,ab.
33. "cognistat".ti,ab.
34. "quick cognitive screening test".ti,ab.
35. "QCST".ti,ab.
36. "rapid dementia screening test".ti,ab.
37. "RDST".ti,ab.
38. "repeatable battery for the assessment of neuropsychological status".ti,ab.
39. "RBANS".ti,ab.
40. "rowland universal dementia assessment scale".ti,ab.
41. "rudas".ti,ab.
42. "self-administered gerocognitive exam*".ti,ab.
43. ("self-administered" and "SAGE").ti,ab.
44. "self-administered computerized screening test for dementia".ti,ab.
45. "short and sweet screening instrument".ti,ab.
46. "sassi".ti,ab.
47. "short cognitive performance test".ti,ab.
48. "syndrome kurztest".ti,ab.
49. "six item screener".ti,ab.
50. "short memory questionnaire".ti,ab.
51. ("short memory questionnaire" and "SMQ").ti,ab.
52. "short orientation memory concentration test".ti,ab.
53. "s-omc".ti,ab.
54. "short blessed test".ti,ab.
55. "short portable mental status questionnaire".ti,ab.
56. "spmsq".ti,ab.
57. "short test of mental status".ti,ab.
58. "telephone interview of cognitive status modified".ti,ab.
59. "tics-m".ti,ab.
60. "trail making test".ti,ab.
61. "verbal fluency categories".ti,ab.
62. "WORLD test".ti,ab.
63. "general practitioner assessment of cognition".ti,ab.
64. "GPCOG".ti,ab.

(Continued)

65. "Hopkins verbal learning test".ti,ab.

66. "HVLT".ti,ab.

67. "time and change test".ti,ab.

68. "modified world test".ti,ab.

69. "symptoms of dementia screener".ti,ab.

70. "dementia questionnaire".ti,ab.

71. "7MS".ti,ab.

72. ("concord informant dementia scale" or CIDS).ti,ab.

73. (SAPH or "dementia screening and perceived harm*").ti,ab.

74. or/1-73

75. exp Dementia/

76. Delirium, Dementia, Amnestic, Cognitive Disorders/ OR Cognition Disorders/ OR Memory Disorders/

77. dement*.ti,ab.

78. alzheimer*.ti,ab.

79. AD.ti,ab.

80. ("lewy bod*" or DLB or LBD).ti,ab.

81. "cognit* impair*".ti,ab.

82. (cognit* adj4 (disorder* or declin* or fail* or function*)).ti,ab.

83. (memory adj3 (complain* or declin* or function*)).ti,ab.

84. or/75-83

85. exp "sensitivity and specificity"/

86. "reproducibility of results"/

87. (predict* adj3 (dement* or AD or alzheimer*)).ti,ab.

88. (identif* adj3 (dement* or AD or alzheimer*)).ti,ab.

89. (discriminat* adj3 (dement* or AD or alzheimer*)).ti,ab.

90. (distinguish* adj3 (dement* or AD or alzheimer*)).ti,ab.

91. (differenti* adj3 (dement* or AD or alzheimer*)).ti,ab.

92. diagnos*.ti.

93. di.fs.

94. sensitivit*.ab.

95. specificit*.ab.

96. (ROC or "receiver operat*").ab.

97. Area under curve/

98. ("Area under curve" or AUC).ab.

(Continued)

99. (detect* adj3 (dement* or AD or alzheimer*)).ti,ab.

100. sROC.ab.

101. accura*.ti,ab.

102. (likelihood adj3 (ratio* or function*)).ab.

103. (conver* adj3 (dement* or AD or alzheimer*)).ti,ab.

104. ((true or false) adj3 (positive* or negative*)).ab.

105. ((positive* or negative* or false or true) adj3 rate*).ti,ab.

106. or/85-105

107. exp dementia/di OR Delirium, Dementia, Amnestic, Cognitive Disorders/di

108. Cognition Disorders/di

109. Memory Disorders/di

110. or/107-109

111. *Neuropsychological Tests/

112. *Questionnaires/

113. Geriatric Assessment/mt

114. *Geriatric Assessment/

115. Neuropsychological Tests/mt, st

116. "neuropsychological test*".ti,ab.

117. (neuropsychological adj (assess* or evaluat* or test*)).ti,ab.

118. (neuropsychological adj (assess* or evaluat* or test* or exam* or battery)).ti,ab.

119. Self report/

120. self-assessment/ or diagnostic self evaluation/

121. Mass Screening/

122. early diagnosis/

123. or/111-122

124. 74 or 123

125. 110 and 124

126. 74 or 123

127. 84 and 106 and 126

128. 74 and 106

129. 125 or 127 or 128

130. (animals not (humans and animals)).sh.

131. 129 not 130

The concepts for this are:

(Continued)

A Specific neuropsychological tests

B General terms (both free text and MeSH) for tests/testing/screening

C Outcome: dementia diagnosis (unfocused MeSH with diagnostic sub-headings)

D Condition of interest: Dementia (general dementia terms both free text and MeSH – exploded and unfocused)

E Methodological filter: not used to limit all search

The concept combinations are:

1. (A OR B) AND C

2. (A OR B) AND D AND E

3. A AND E

Appendix 7. Assessment of reporting quality - STARD checklist

Section and Topic		
TITLE/ABSTRACT KEYWORDS	1	Identify the article as a study of diagnostic accuracy (recommend MeSH heading 'sensitivity and specificity').
INTRODUCTION	2	State the research questions or study aims, such as estimating diagnostic accuracy or comparing accuracy between tests or across participant groups.
METHODS		
Participants	3	The study population: The inclusion and exclusion criteria, setting and locations where data were collected.
	4	Participant recruitment: Was recruitment based on presenting symptoms, results from previous tests, or the fact that the participants had received the index tests or the reference standard?
	5	Participant sampling: Was the study population a consecutive series of participants defined by the selection criteria in item 3 and 4? If not, specify how participants were further selected.
	6	Data collection: Was data collection planned before the index test and reference standard were performed (prospective study) or after (retrospective study)?
Test methods	7	The reference standard and its rationale.
	8	Technical specifications of material and methods involved including how and when measurements were taken, and/or cite references for index tests and reference standard.
	9	Definition of and rationale for the units, cut-offs and/or categories of the results of the index tests and the reference standard.

(Continued)

	10	The number, training and expertise of the persons executing and reading the index tests and the reference standard.
	11	Whether or not the readers of the index tests and reference standard were blind (masked) to the results of the other test and describe any other clinical information available to the readers.
Statistical methods	12	Methods for calculating or comparing measures of diagnostic accuracy, and the statistical methods used to quantify uncertainty (e.g. 95% confidence intervals).
	13	Methods for calculating test reproducibility, if done.
RESULTS		
Participants	14	When study was performed, including beginning and end dates of recruitment.
	15	Clinical and demographic characteristics of the study population (at least information on age, gender, spectrum of presenting symptoms).
	16	The number of participants satisfying the criteria for inclusion who did or did not undergo the index tests and/or the reference standard; describe why participants failed to undergo either test (a flow diagram is strongly recommended).
Test results	17	Time-interval between the index tests and the reference standard, and any treatment administered in between.
	18	Distribution of severity of disease (define criteria) in those with the target condition; other diagnoses in participants without the target condition.
	19	A cross tabulation of the results of the index tests (including indeterminate and missing results) by the results of the reference standard; for continuous results, the distribution of the test results by the results of the reference standard.
	20	Any adverse events from performing the index tests or the reference standard.
Estimates	21	Estimates of diagnostic accuracy and measures of statistical uncertainty (e.g. 95% confidence intervals).
	22	How indeterminate results, missing data and outliers of the index tests were handled.
	23	Estimates of variability of diagnostic accuracy between subgroups of participants, readers or centres, if done.
	24	Estimates of test reproducibility, if done.
DISCUSSION	25	Discuss the clinical applicability of the study findings.

Appendix 8. Assessment of methodological quality table QUADAS-2 tool

DOMAIN	PATIENT SELECTION	INDEX TEST	REFERENCE STANDARD	FLOW AND TIMING

(Continued)

Description	Describe methods of patient selection: Describe included patients (prior testing, presentation, intended use of index test and setting):	Describe the index test and how it was conducted and interpreted:	Describe the reference standard and how it was conducted and interpreted:	Describe any patients who did not receive the index test(s) and/or reference standard or who were excluded from the 2x2 table (refer to flow diagram): Describe the time interval and any interventions between index test(s) and reference standard:
Signalling questions (yes/no/unclear)	Was a consecutive or random sample of patients enrolled?	Were the index test results interpreted without knowledge of the results of the reference standard?	Is the reference standard likely to correctly classify the target condition?	Was there an appropriate interval between index test(s) and reference standard?
	Was a case-control design avoided?	If a threshold was used, was it pre-specified?	Were the reference standard results interpreted without knowledge of the results of the index test?	Did all patients receive a reference standard?
	Did the study avoid inappropriate exclusions?			Did all patients receive the same reference standard?
				Were all patients included in the analysis?
Risk of bias: High/ low/ unclear	Could the selection of patients have introduced bias?	Could the conduct or interpretation of the index test have introduced bias?	Could the reference standard, its conduct, or its interpretation have introduced bias?	Could the patient flow have introduced bias?
Concerns regarding applicability: High/ low/ unclear	Are there concerns that the included patients do not match the review question?	Are there concerns that the index test, its conduct, or interpretation differ from the review question?	Are there concerns that the target condition as defined by the reference standard does not match the review question?	

Appendix 9. Anchoring statements for quality assessment of AD-8 diagnostic studies

We provide some core anchoring statements for quality assessment of diagnostic test accuracy reviews of AD-8 in dementia. These statements are designed for use with the QUADAS-2 tool and were derived during a two-day, multidisciplinary focus group.Davis 2013

During the focus group and the piloting/validation of this guidance, it was clear that certain issues were key to assessing quality, while other issues were important to record but less important for assessing overall quality. To assist, we describe a system wherein certain items can dominate. For these dominant items, if scored “high risk” then that section of the QUADAS-2 results table is likely to be scored as high risk

of bias regardless of other scores. For example, in dementia diagnostic test accuracy studies, ensuring that clinicians performing dementia assessment are blinded to results of index test is fundamental. If this blinding was not present then the item on reference standard should be scored “high risk of bias”, regardless of the other contributory elements.

We have detailed how QUADAS-2 has been operationalised for use with dementia reference standard studies below. In these descriptors dominant items are labelled as "high risk of bias for total section regardless of other items".

In assessing individual items, the score of unclear should only be given if there is genuine uncertainty. In these situations review authors will contact the relevant study teams for additional information.

Anchoring statements to assist with assessment for risk of bias

Selection

Was a case-control or similar design avoided?

Designs similar to case-control that may introduce bias are those designs where the study team deliberately increase or decrease the proportion with the target condition. For example, a population study may be enriched with extra dementia patients from a secondary care setting. Such studies will be automatically labelled high risk of bias and this will be assessed as a potential source of heterogeneity.

If case-control is used then grading will be high risk of bias for total section regardless of other items (in fact case-control studies will not be included in this review).

Was the sampling method appropriate?

Where sampling is used, the designs least likely to cause bias are consecutive sampling or random sampling. Sampling that is based on volunteers or selecting participants from a clinic or research resource is prone to bias.

Are exclusion criteria described and appropriate?

The study will be automatically graded as unclear if exclusions are not detailed (pending contact with study authors). Where exclusions are detailed, the study will be graded as low risk of bias if exclusions are felt to be appropriate by the review authors. Certain exclusions common to many studies of dementia are: medical instability; terminal disease; alcohol/substance misuse; concomitant psychiatric diagnosis; other neurodegenerative condition.

Post hoc exclusions will be labelled high risk of bias for total section regardless of other items.

Index Test

Was AD-8 assessment performed without knowledge of clinical dementia diagnosis?

Terms such as “blinded” or “independently and without knowledge of” are sufficient and full details of the blinding procedure are not required. This item may be scored as low risk of bias if explicitly described or if there is a clear temporal pattern to order of testing that precludes the need for formal blinding i.e. all AD-8 assessments performed before dementia assessment.

If there is no attempt at blinding grading will be high risk of bias for total section regardless of other items.

Were AD-8 thresholds prespecified?

For scales there is often a reference point (in units or categories) above which participants are classified as “test positive”; this may be referred to as threshold; clinical cut-off or dichotomisation point. A study is classified high risk of bias if the authors define the optimal cut-off post-hoc based on their own study data. Certain papers may use an alternative methodology for analysis that does not use thresholds and these papers should be classified as low risk of bias.

Were sufficient data on AD-8 application given for the test to be repeated in an independent study?

Particular points of interest for AD-8 include method of administration (for example, self-completed questionnaire versus direct questioning interview); nature of informant; language of assessment. If a novel form of AD-8 is used, details of the scale should be included or a reference given to an appropriate descriptive text. Where AD-8 is used in a novel manner, for example, a translated questionnaire, there should be evidence of validation work.

Reference Standard

Is the assessment used for clinical diagnosis of dementia acceptable?

Commonly used international criteria to assist with clinical diagnosis of dementia include those detailed in DSM-IV and ICD-10. Criteria specific to dementia subtypes include but are not limited to NINCDS-ADRDA criteria for Alzheimer’s dementia; McKeith criteria for Lewy Body dementia; Lund criteria for frontotemporal dementias; and the NINDS-AIREN criteria for vascular dementia. Where the criteria used

for assessment are not familiar to the review authors or the Cochrane Dementia and Cognitive Improvement Group this item should be classified as high risk of bias.

Was clinical assessment for dementia performed without knowledge of AD-8?

Terms such as "blinded" or "independent" are sufficient and full details of the blinding procedure are not required. This may be scored as low risk of bias if explicitly described or if there is a clear temporal pattern to order of testing, i.e. all dementia assessments performed before AD-8 testing.

Informant rating scales and direct cognitive tests present certain problems. It is accepted that informant interview and cognitive testing is a usual component of clinical assessment for dementia, however, specific use of the scale under review in the clinical dementia assessment should be scored as high risk of bias. We have prespecified that dementia diagnosis that explicitly uses AD-8 will be classified as high risk of bias for total section regardless of other items.

Were sufficient data on dementia assessment method given for the assessment to be repeated in an independent study?

The criteria used for clinical assessment are discussed in another item. Particular points of interest for dementia assessment include the background of the assessor, training/expertise of the assessor; additional information available to inform diagnosis (neuroimaging; neuropsychological testing).

Flow

Was there an appropriate interval between AD-8 and clinical dementia assessment?

For a cross-sectional study design, there is potential for change between assessments. The ideal would be same day assessment but this is not always feasible. We have set an arbitrary maximum interval of one month between tests, although this may be revised depending on the test and the stability of the condition of interest.

For a study looking at delayed verification there is no agreement on how long the interval should be between index test and first/last assessment for dementia. An interval of less than six months is unlikely to be sufficient time for progression.

Did all get the same assessment for dementia regardless of AD-8 result?

There may be scenarios where only those who score "test positive" on AD-8 have a more detailed assessment. Where dementia assessment (or other reference standard) differs depending on the AD-8 result this should be classified as high risk of bias.

Were all who received AD-8 assessment included in the final analysis?

If the study has drop outs these should be accounted for; a maximum proportion of drop outs to remain low risk of bias has been specified as 20%.

Were missing AD-8 results or un-interpretable AD-8 results reported?

Where missing results are reported if there is substantial attrition (we have set an arbitrary value of 50% missing data) this should be scored as high risk of bias for total section regardless of other items.

Applicability

Were those included representative of the general population of interest?

Those included should match the intended population as described in the review question. If not already specified in the review inclusion criteria, setting will be particularly important – the review authors should consider population in terms of symptoms; pre-testing; potential disease prevalence. Studies that use very selected groups or subgroups will be classified as poor applicability.

Was AD-8 performed consistently and in a manner similar to its use in clinical practice?

AD-8 studies will be judged against the original description of its use.

Was clinical diagnosis of dementia (or other reference standard) made in a manner similar to current clinical practice?

For many reviews, inclusion criteria and assessment for risk of bias will already have assessed the dementia diagnosis. For certain reviews an applicability statement relating to reference standard may not be applicable. There is the possibility that a form of dementia assessment, although valid, may diagnose a far larger proportion with disease than would be seen in usual clinical practice. In this instance the item should be rated poor applicability.

DECLARATIONS OF INTEREST

No authors have any relevant disclosures of conflicts of interest.